

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 26, 2003, 11:15:29 ; Search time 1171 Seconds
(without alignments)
30.445 Million cell updates/sec

Title: US-10-086-184-2

Perfect score: 21

Sequence: 1 gtcctactgacagagctgacc 21

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 60474

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*

1: em_escba:*
2: em_escba:*
3: em_escba:*
4: em_escba:*
5: em_escba:*
6: em_escba:*
7: em_escba:*
8: em_escba:*
9: em_escba:*
10: em_escba:*
11: em_escba:*
12: em_escba:*
13: em_escba:*
14: em_escba:*
15: em_escba:*
16: em_escba:*
17: em_escba:*
18: em_escba:*
19: em_escba:*
20: em_escba:*
21: em_escba:*
22: em_escba:*
23: em_escba:*
24: em_escba:*
25: em_escba:*
26: em_escba:*
27: em_escba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12.2	55.1	35	17	AZ819702 2M0091A01
2	12	57.1	39	17	BH792327 SALK 0634
3	11.8	56.2	30	17	AZ780490 2M0017524
4	11.6	55.2	33	17	AL743211 Danilo rex
5	11.6	55.2	34	17	BH812254 SALK 0615
6	11.6	55.2	37	17	AZ465387 1M0275113

7	11.4	54.3	29	17	AZ794355
8	11.4	54.3	37	17	AZ507623
9	11.4	54.3	38	17	AZ783438
10	11.4	54.3	38	17	BH789778
11	11.4	54.3	39	17	BH855660
12	11.2	53.3	37	9	AA444863
13	11.2	53.3	37	17	BH850593
14	11	52.4	32	17	AZ496170
15	11	52.4	31	14	R68337
16	11	52.4	37	17	AO073641
17	11	52.4	40	17	AZ363389
18	11	52.4	40	17	BH759445
19	10.8	51.4	22	17	TA235A09Q
20	10.8	51.4	22	17	TA334H05Q
21	10.8	51.4	27	17	AZ439712
22	10.8	51.4	27	17	TA130G05P
23	10.8	51.4	32	13	BJ065561
24	10.6	50.5	32	17	AZ407990
25	10.6	50.5	34	17	AO025206
26	10.6	50.5	36	17	TA119B10Q
27	10.6	50.5	37	9	AA877091
28	10.6	50.5	37	17	AZ458505
29	10.6	50.5	40	17	BH800126
30	10.4	49.5	22	17	AZ806828
31	10.4	49.5	26	17	AZ619253
32	10.4	49.5	26	17	TA120A05P
33	10.4	49.5	31	17	AZ856886
34	10.4	49.5	33	17	AZ451738
35	10.4	49.5	33	17	AZ481982
36	10.4	49.5	35	17	BH811059
37	10.4	49.5	35	17	BH851755
38	10.4	49.5	38	10	AV969052
39	10.4	49.5	39	12	BE874819
40	10.4	49.5	40	9	AI023350
41	10.4	49.5	40	9	AI666881
42	10.2	48.6	26	17	TA336D06Q
43	10.2	48.6	27	17	BH791867
44	10.2	48.6	30	17	AZ941284
45	10.2	48.6	31	9	AI016099

ALIGNMENTS

RESULT 1
LOCUS AZ819702 35 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0091A01R Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC2M0091A01 R, DNA sequence.
ACCESSION AZ819702
VERSION AZ819702.1 GI:12989610
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 35)
AUTHORS Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A.
and Wright,D., Weis,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0091 row: A column: 01
Seq primer: CACACAGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 35.
Location/Qualifiers

FEATURES

source

1.35
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0091A01"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|g14732072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E.coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

10 a 6 c 4 g 15 t

ORIGIN

Query Match

58.1%; Score 12.2; DB 17; Length 35;

Best Local Similarity

82.4%; Pred. No. 3.6e+04; Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY

2 TGCTACTGATGAGTGT 18

Db

26 TGTTCAGAAAGACTGT 10

RESULT 2

BH792327

LOCUS

39 bp DNA linear GSS 02-APR-2002
SALK_063427.45.40.x Arabidopsis thaliana TDNA insertion lines

DEFINITION

Arabidopsis thaliana genomic clone SALK_063427.45.40.x, DNA sequence.

ACCESSION

BH792327

VERSION

BH792327

KEYWORDS

GSS.

SOURCE

Chale cress

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadriab,C., Jecker,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J., and Ecker,J.R.

AUTHORS

A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

TITLE

Unpublished (2001)

JOURNAL

Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

TDNA.
Class: TDNA tagged
Location/Qualifiers

FEATURES

source

1.39
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_063427.45.40.x"
/clone_1lb="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT

11 a 6 c 6 g 16 t

ORIGIN

Query Match

57.1%; Score 12; DB 17; Length 39;

Best Local Similarity

75.0%; Pred. No. 4.8e+04; Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY

2 TGCTACTGATGAGTGTACC 21

Db

9 TTCTTCGATGATGACACC 28

RESULT 3

AZ780490/c

LOCUS

30 bp DNA linear GSS 16-FEB-2001
2M0017G24R Mouse 10kb plasmid UUCG1M library Mus musculus genomic

DEFINITION

Clone UUCG2M0017G24 R, DNA sequence.

ACCESSION

AZ780490

VERSION

AZ780490

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,D., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weis,R.

AUTHORS

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE

Unpublished (2000)

JOURNAL

Contact: Robert B. Weis
University of Utah Genome Center
University of Utah

COMMENT

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

FEATURES

Insert length: 10000 Std Error: 0.00
Plate: 0017 row: G column: 24
Seq primer: CACACAGAAACGCTATGACC
Class: plasmid ends
High quality sequence stop: 30.
Location/Qualifiers

source

1.30
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0017G24"
/clone_1lb="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (GI4732114|gb|AF12907.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

9 a 10 c 3 g 8 t

ORIGIN

Query Match

Best Local Similarity 56.2%; Score 11.8; DB 17; Length 30;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 5 TACTGATAGAGTGA 19

Db 18 TACAATAGAGTGA 4

RESULT 4

DR19023S

LOCUS DR19023S 33 bp DNA linear GSS 06-JUN-2002

DEFINITION Dario rerio genomic clone DKey-19023, genomic survey sequence.

ACCESSION AL743211

VERSION AL743211.1 GI:21346185

KEYWORDS GSS;

SOURCE zebrafish.

ORGANISM Dario rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 33)

Humphray,S.J., Huckle,E. and Hunt,S.E.

TITLE Direct Submission

JOURNAL Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:

humphrey@sanger.ac.uk Unpublished

COMMENT This sequence was generated from the SP6 end of BAC 19023. 19023 is part of the Dariokey Pilot BAC Library created by R. Plasterk and N.V. Keygene.

Further details: http://www.sanger.ac.uk/Projects/D_rerio/.

location/Qualifiers

1..33

/organism="Dario rerio"

/db_xref="taxon:7955"

/clone="DKey-19023"

/issue_type="Testis"

/note="Vector pIndigoBAC-536"

BASE COUNT 10 a 6 c 7 g 10 t

ORIGIN

Query Match

Best Local Similarity 55.2%; Score 11.6; DB 17; Length 33;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 TCGTACTGATAGAGTGA 19

Db 10 TCGTACTGATAGAGTGA 27

RESULT 5

BH812254/c

LOCUS BH812254/c

DEFINITION SALK 061508 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_061508, DNA sequence.

ACCESSION

BH812254

VERSION BH812254.1 GI:20390709

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosida II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 34)

Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab

,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

, Zimmermann,J., and Ecker,U.R.

A Sequence-Indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished (2001)

CONTACT: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA. This sequence lies within 300 bases of the 3' end of

At5g40710.

Class: TDNA tagged.

location/Qualifiers

1..34

/organism="Arabidopsis thaliana"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="SALK_061508"

/note="PCR was performed on Arabidopsis thaliana TDNA insertion

elements. The resultant fragment for each line was

directly sequenced to determine the genomic sequence at

the site of insertion. Details of the protocols used can

be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 9 a 11 c 5 g 9 t

ORIGIN

Query Match

Best Local Similarity 55.2%; Score 11.6; DB 17; Length 34;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 3 GCTACTGATAGAGTGA 20

Db 18 GCGGCTGAGAGAGATAC 1

RESULT 6

AZ465387

LOCUS AZ465387

DEFINITION 1M0275113F Mouse 10kb Plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0275113 F, DNA sequence.

ACCESSION AZ465387

VERSION AZ465387.1 GI:10623512

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 37)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

Plasmid inserts

Unpublished (2000)

CONTACT: Robert B. Weiss

University of Utah Genome Center

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0275 row: 1 column: 13
Seq primer: CGTGTAAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 37.
Location/Qualifiers

FEATURES

source

1..37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0275113"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (G1|4732114|9b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

9 a 1 c 19 g 8 t

ORIGIN

Query Match 55.2%; Score 11.6; DB 17; Length 37;
Best Local Similarity 77.8%; Pred. No. 7.3e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GTGCTACTGATGAGTGT 18
Db 19 GGCTAGTGTAGAGAGT 36

RESULT 7

A2794355

LOCUS

29 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0048A1F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

A2794355

ACCESSION

A2794355

VERSION

A2794355.1

GSS.

KEYWORDS

house mouse.

SOURCE

ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 29)
REFERENCE
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0048 row: A column: 11
Seq primer: CGTGTAAACGACGCGCACT
Class: plasmid ends
High quality sequence stop: 29.
Location/Qualifiers

FEATURES

source

1..29
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0048A11"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (G1|4732114|9b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

5 a 7 c 10 g 7 t

ORIGIN

Query Match 54.3%; Score 11.4; DB 17; Length 29;
Best Local Similarity 71.4%; Pred. No. 8.2e+04;
Matches 15; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 GTGCTACTGATGAGTGTACC 21
Db 7 GTGCAATGATGATGATGAC 27

RESULT 8

A2507623

LOCUS

37 bp DNA linear GSS 05-OCT-2000
DEFINITION 1M0349A20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

A2507623

ACCESSION

A2507623

VERSION

A2507623.1

GSS.

KEYWORDS

house mouse.

SOURCE

ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 37)
REFERENCE
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weis, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0349 row: A column: 20
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 37.

FEATURES

source

1. .37
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0349420"
/clone_11b="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

4 a 2 c 12 g 19 t

ORIGIN

Query Match 54.3%; Score 11.4; DB 17; Length 37;
Best Local Similarity 71.4%; Pred. No. 9.2e+04;
Matches 15; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GTGCTACTGATAGTGATACC 21
Db 2 GTGTTGTGTTGAGTGATAC 22

RESULT 9
A2783438 38 bp DNA linear GSS 16-FEB-2001
LOCUS 2M0025H12F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
DEFINITION clone UUCGCM0025H12 F, DNA sequence.
ACCESSION A2783438
VERSION A2783438.1 GI:12918166
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 38)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0025 row: H column: 12
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 38.

FEATURES

source

1. .38
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0025H12"
/clone_11b="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

10 a 10 c 11 g 7 t

ORIGIN

Query Match 54.3%; Score 11.4; DB 17; Length 38;
Best Local Similarity 71.4%; Pred. No. 9.3e+04;
Matches 15; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 GTGCTACTGATAGTGATACC 21
Db 1 GTGCTAAGGTGAGAGAAATAC 21

RESULT 10
BH789778 38 bp DNA linear GSS 02-APR-2002
LOCUS SALK_046680.48.35.x Arabidopsis thaliana TDNA insertion line
DEFINITION Arabidopsis thaliana genomic clone SALK_046680.48.35.x, DNA sequence.
ACCESSION BH789778
VERSION BH789778.1 GI:19882876
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 38)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadgilnab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shim, P., Zimmermann, J. and Ecker, J.R.
TITLE A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated exon of At1g51260.
Class: TDNA tagged.

FEATURES
source
Location/Qualifiers

1. 38
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_046680.48.35.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"

BASE COUNT

8 a 9 c 10 g 11 t

Query Match 54.3%; Score 11.4; DB 17; Length 39;
Best Local Similarity 92.3%; Pred. No. 9.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3. GCCTACTGATGAG 15
Db 1 GCCTTGATGAG 13

RESULT 11

BH855660

LOCUS

DEFINITION BH855660 39 bp DNA linear GSS 08-JUL-2002

Arabidopsis thaliana genomic clone SALK_084826.49.95.x, DNA

ACCESSION

VERSION BH855660

KEYWORDS

SOURCE

ORGANISM

Chale crees.
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 39)

Alonso,J.M., Leisbe,T.J., Barajas,P., Chen,H., Cheuk,R., Gadri,nab

, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Predits,L., Shinn,P.

, Zimmerman,J. and Ecker,J.R.

A Sequence-indexed Library of Insertion Mutations in the

Arabidopsis Genome

Unpublished (2001)

Contact: Joseph R. Ecker

Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies

10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Tel: 858 453 4100 x1752

Fax: 858 558 6379

Email: ecker@salk.edu

This is single pass sequence recovered from the left border of

TDNA.

Class: TDNA tagged.

Location/Qualifiers

1. 39
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_084826.49.95.x"
/note="PCR was performed on Arabidopsis thaliana lines

FEATURES

source

BASE COUNT

each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"

Query Match 54.3%; Score 11.4; DB 17; Length 39;
Best Local Similarity 85.7%; Pred. No. 9.5e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TGCTACTGATGAG 15
Db 2 TGCTACTGATGAG 15

RESULT 12

AA444863/c

LOCUS

DEFINITION AA444863 37 bp mRNA linear EST 03-JUN-1997

ve75f02.r1 Soares_mammary_gland_NBMKG Mus musculus cDNA clone

IMAGE:832059 5' similar to gb:U70423 M.musculus mRNA for monoclonal

antibody heavy chain gamma (MOUSE);, mRNA sequence.

AA444863.1 GI:2157253

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 37)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Giesel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Jan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LBNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:492275

Trace considered overall poor quality

Seq primer: -28m3 rev2 ET from Amersham

High quality sequence step: 1.

Location/Qualifiers

1. 37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:832059"
/clone_lib="Soares_mammary_gland_NBMKG"
/sex="male"
/issue_type="mammary gland"
/dev stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pRT73D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TGTTCACATCTGAGTGGAGCGCGCGAATGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pRT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT	8 a	14 c	8 g	7 t
ORIGIN				
Query Match	53.3%	Score 11.2;	DB 9;	Length 37;
Best Local Similarity	81.2%;	Pred. No. 1.2e+05;		
Matches 13; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;
Oy	6	ACTGATAGCTGACC 21		
Db	29	ACTGAGAGGTGACC 14		

RESULT 13

LOCUS	BH850593	37 bp	DNA	linear	GSS 13-JUN-2002
DEFINITION	SALK_071553.14.35.x Arabidopsis thaliana TDNA insertion lines				
ACCESSION	BH850593				
VERSION	BH850593				
KEYWORDS	BH850593.1	GI:21421464			
SOURCE	GSS.				
ORGANISM	thale cress.				
REFERENCE	Arabidopsis thaliana				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 37)				
TITLE	Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R				
JOURNAL	A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome				
COMMENT	Unpublished (2001)				
	Contact: Joseph R. Ecker				
	Salk Institute Genomic Analysis Laboratory (Signal)				
	The Salk Institute for Biological Studies				
	10010 N. Torrey Pines Road, La Jolla, CA 92037, USA				
	Tel.: 858 453 4100 x1752				
	Fax: 858 558 6379				
	Email: eckersalk.edu				
	This is single pass sequence recovered from the left border of TDNA.				
FEATURES	Class: TDNA tagged.				
source	Location/Qualifiers				
	1..37				
	/organism="Arabidopsis thaliana"				
	/strain="Columbia 0"				
	/db_xref="taxon:3702"				
	/clone="SALK_071553.14.35.x"				
	/clone_1b="Arabidopsis thaliana TDNA insertion lines"				
	/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html "				
BASE COUNT	16 a	2 c	10 g	9 t	
ORIGIN					
Query Match	53.3%	Score 11.2;	DB 17;	Length 37;	
Best Local Similarity	81.2%;	Pred. No. 1.2e+05;			
Matches 13; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;	
Oy	6	ACTGATAGCTGACC 21			
Db	22	ACTGATAGCTGACC 37			

RESULT 14

LOCUS	AZ496170	22 bp	DNA	linear	GSS 05-OCT-2000
DEFINITION	IM0332N13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0332N13 F, DNA sequence.				

ACCESSION	AZ496170
VERSION	AZ496170.1
KEYWORDS	GI:10671985
SOURCE	GS.
ORGANISM	house mouse.
REFERENCE	Mus musculus. Mammalia; Eutheria; Rodentia; Sclerognathis; Muridae; Murinae; Mus. 1 (bases 1 to 22)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly, 'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weis,R. Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts Unpublished (2000)
JOURNAL	Contact: Robert B. Weis University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel.: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert length: 10000 Std Error: 0.00 Plate: 0332 row: N column: 13 Seq primer: CGTTGTAAAACGACGGCCAGT Class: plasmid ends High quality sequence stop: 22. Location/Qualifiers
FEATURES	1. 22 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC1M0332N13" /clone_lib="Mouse 10kb plasmid UUGC1M library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g11473114[gb]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT	7 a 4 c 5 g 6 t
ORIGIN	
Query Match	52.4%; Score 11; DB 17; Length 22;
Best Local Similarity	73.7%; Pred. No. 1.1e+05;
Matches 14; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
0y	2 TGCTACTGATAGAGTGATAC 20 1 RG8337/c
Db	1 TGATATATATGCACTGTAC 19
RESULT 15	
LOCUS	RG8337 31 bp mRNA linear EST 01-JUN-1995
DEFINITION	Y106H02.r1 Soares Placenta NB2HP Homo sapiens cDNA clone IMAGE:138483 5' similar to SP:S33957 S33957; COAT PROTEIN GAMMA-COP

```

- BOVINE ;, mRNA sequence.
ACCESSION   R68337
VERSION     R68337.1
KEYWORDS    GI:841854
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 31)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Woldmann,P. and Wilson,R.
            The WashU-Merck EST Project
TITLE       Unpublished (1995)
JOURNAL     Contact: Wilson RK
COMMENT     Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.wustl.edu
            Insert Size: 1412
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LBNL
            This clone is available royalty-free through LBNL; contact the
            IMAGE Consortium (info@image.lbnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert length: 1412 Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES
    source
        1..31
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="GDB:544867"
                /db_xref="taxon:9606"
                /clone="IMAGE:138483"
                /clone_1b="Soares placenta Nb2HP"
                /sex="Female"
                /dev_stage="Placenta obtained at birth (full term)"
                /lab_host="DH10B (ampicillin resistant)"
                /note="Organ: Placenta; Vector: pT73D (Pharmacia) with a
                modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                strand cDNA was primed with a Not I - oligo(dT) primer (5'
                AACTGAGAGATTCCGCGCCGACAGATTATTTTATTTT 3'),
                double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Not I and cloned into the Not I
                and Eco RI sites of the modified pT73 vector. Library
                went through one round of normalization. Library
                constructed by Bento Soares and M.Fatima Bonaldo. "
BASE COUNT   13 a          9 c          4 g          5 t
ORIGIN
Query Match 52.4%; Score 11; DB 14; Length 31;
Best Local Similarity 73.7%; Pred. No. 1.3e+05;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 TGCTACTGATAGAGGTAC 20
DB 22 TGATGTTCTAGAGGTAC 4

```

Search completed: March 26, 2003, 23:18:00
 Job time : 11173 secs